Abstract: In this paper, we demonstrate that Brazilian Portuguese phonemic nasal vowels are, as Mattoso Camara (1953, 1970) suggested, derived from an underlying sequence of oral vowel plus nasal consonant... or almost that. Two types of transformational language games reveal that Mattoso was 80% right: four of the nasal vowels are indeed derived from such a sequence, while nasal [ã] (strictly speaking, [ê]) is the result of an underlying, inherently nasal vowel. We also propose a further revision to Mattoso’s disegmental hypothesis for nasal vowels: rather than being a sequence of oral vowel plus nasal archipheneme, we propose that they are specifically the result of a combination of oral vowel plus homorganic nasal glide (i.e. [j] or [w]). This nasal glide, however, is incompatible with /ã/, and moreover would disrupt the contrast between [ẽ], [ê], and [êw]. As a result, [ẽ] must be inherently underlyingly nasal, a result confirmed through a variety of external evidence.

Keywords: phonemic nasal vowels, invented language games, Catete language game

Resumo: Neste artigo, demonstramos que as vogais nasais do Português do Brasil são, como propôs Mattoso Câmara (1953, 1970), o resultado superficial de uma sequência de vogal oral e consoante nasal... ou quase isso. Dois jogos de linguagem, com diferentes tipos de transformação, revelam que Mattoso tinha 80% de razão: quatro das cinco vogais nasais são, de fato, resultantes de uma sequência V+N, enquanto [ã] (mais precisamente, [ê]) seria a realização de uma vogal subjetivamente nasal. Além disso, propomos outra revisão da hipótese Mattosiana: para as vogais nasais bisegmentais, em vez...
1. Mattoso’s problem and transformational language games

While Plato’s Problem as applied to language acquisition is well-known (Chomsky 1986), in this paper we focus on what we call Mattoso’s Problem, in honor of Mattoso Câmara Jr. Mattoso’s problem can be summarized as “What elements of representation should be taken to be atomic, and what elements should be taken to be derived?” The specific question at hand relates to nasal vowels in Portuguese, and whether they should be taken to be basic underlying elements of the inventory, or derived by phonological processes (basically, assimilation of the [+nasal] feature from an underlying nasal segment in coda position that is then subsequently deleted). Arguments that definitively decide between the monosegmental vs. disegmental analysis of nasal vowels can be somewhat hard to come by, as we will shortly see. There are, however, potential ways out of this labyrinth, one being the use of language games as a tool to reveal hidden phonotactic structure.

Other questions of the sort “at what level of representation is element E introduced?” resonate in different areas of the grammar. Take wh- question elements, for which a wide variety of research (wanna-contraction, pronunciation of lower copies, T-to-C movement in embedded clauses, etc) suggests that their surface position is not equal to their underlying position(s). Transformational rules of this sort, and their interaction with other processes, can thus provide evidence for “when” and “where” in the grammar certain surface elements may have been introduced.

In phonology, the range of transformations observable ‘on a daily basis’ is undeniably more limited than those in syntax. However, numerous phonological transformations exist, both ‘in the wild’ and in the laboratory.

---

For a comprehensive overview of the literature, see Moraes & Wetzels (1992), who provide a more fine-grained typology of the analyses for nasal vowels in Brazilian Portuguese. For in depth descriptions of Brazilian Portuguese nasal vowels at the phonetic level, see Souza (1994), Seara (2000), Shosted (2003) and Mendes (2008).
Our focus in this paper, therefore, is on transformational language games that add, delete, replace, or invert segments or syllables, with no apparent trigger from inflectional or derivational morphology. Such games (often called *ludlings*; see Bagemihl (1995)) represent ‘pure transformations’ that can alter or disrupt phonological sequences in just the right way to allow inspection of otherwise undetected moments in the phonological derivation.

The use of transformational language games to explore phonotactic representations enjoys a long history in generative grammar. Chomsky & Halle (1968) look at the output of words such as *ice* vs *sigh* in Pig Latin in order to investigate the productivity of Canadian Raising. Rizzolo (2007) investigates the Serbo-Croatian language game Satrovacki in order to demonstrate that words like *rad* may have an underlying final vowel. In the present work, we bring existing and invented language games to bear on the question of nasal vowels in Brazilian Portuguese.

2. The inventory of nasal and oral vowels in Brazilian Portuguese

2.1. The 7-oral plus 5-nasal vowel system

Brazilian Portuguese has seven distinctive oral vowels [i, e, ɛ, a, ɔ, o, u], and five distinctive nasal vowels [i, ɛ, ð, ʊ, ū], corresponding to a proper subset of the oral vowels (there are also diphthongs, to which we will return). In stressed positions, all of these are distinctive, leading to minimal pairs. Out of all of the contrasts that can be found among pairs of all these twelve vowels, let us focus on the contrasts among nasal vowels, by looking at the minimal pairs in (1), presented for now with a loose phonetic transcription, to be refined later.

(1) **Contrast between nasal vowels** (loose transcription)

| a: [i] : [ɛ] | [sɨ] ‘yes’: [sɛ] ‘without’ |
| b: [i] : [ð] | [fɨ] ‘end’: [fð] ‘fan’ |
| c: [i] : [ʊ] | [ku.ˈpɨ] ‘termite’: [ku.ˈpõ] ‘cupon’ |
| d: [ɨ] : [ũ] | [fiũ] ‘kidney’: [fiũ] ‘rum’ |
| e: [ɛ] : [ð] | [sɛ] ‘without’: [ð] ‘sane(<sub>fem</sub>)’ |
| f: [ɛ] : [o] | [bê] ‘well’: [bõ] ‘good’ |
| g: [ɛ] : [ũ] | [ẽ] ‘in’: [ũ] ‘one’ |
| h: [ð] : [ɔ] | [sð] ‘sane(<sub>fem</sub>)’: [sõ] ‘sound’ |
| i: [ð] : [ũ] | [fið] ‘frog’: [fiũ] ‘rum’ |
| j: [õ] : [u] | [bõ] ‘good’: [bũ] ‘boom’ |
In the following subsections, we will contrast two families of analysis for the twelve surface vowels in question. Before doing so, however, we must clarify that nasal vowels at the phonetic level may be either phonemic or allophonic at the phonological level. For instance, the word [fən] ‘fan’ forms a minimal pair with the word [fa] ‘musical note F’. In this case, we say that [ə] is phonemic, regardless of whether it is a single underlyingly nasal phoneme or the result of a nasalization process involving an abstract nasal coda that later deletes. Allophonic nasal vowels, on the other hand, are predicted to be nasal given the environment in which they appear, and they never contrast with their oral versions. For instance, the nasality on the vowel in the word [fə.əɾ] ‘fame’ is predicted by the presence of the immediately following segment /m/, which involves a nasalization process largely identical to the one required for the word [fən] ‘fan’ under the disegmental analysis, except that the segment spreading its nasality to the vowel is in the onset of the following syllable, rather than in the coda of the same syllable. What is relevant here is that, in nasal vowels of words like [fə.əɾ] ‘fame’, there is never a possible minimal pair exhibiting contrast between a nasal and a non-nasal vowel. That is, words like *[fa.əɾ] are impossible in the language.

Under most disegmental analyses, these phonemic and allophonic nasal vowels are quite similar, though not fully identical. Our focus in this paper is largely on phonemic nasal vowels, discussed in the experiments in section 3. In section 5, we introduce language game data regarding allophonic nasal vowels in order to broaden the discussion.

2.2. Nasal vowels as monosegmental units

The monosegmental analysis takes all twelve vowels to be each a distinct phoneme, meaning that the five surface nasal vowels [ɨ, ɐ, ɐ̃, ɔ, ū] come from underlying forms /ɪ, ɐ, ɐ̃, ɔ, ū/, respectively (cf. Sten 1944, Lüdtke 1953, Stevens 1954, Hammarström 1962, Head 1965, Leite 1974, Mata Machado 1981). Proponents of this analysis argue that evidence for it comes from the existence of minimal pairs between each nasal vowel and a corresponding vowel that is identical to it except for nasality, as in (2).

(2) a: [ɨ] : [ɨ]  [ˈvi.də] ‘life’ : [ˈvi.də] ‘arrival’
b: [ɐ] : [ɐ̃]  [ˈte.tə] ‘tit’ : [ˈtə.ə] ‘(s)he tries’
c: [a] : [ɐ]  [fa] ‘musical note F’ : [fə] ‘fan’
d: [ɔ] : [ɔ̃]  [ˈbo.əɾ] ‘silly’ : [ˈbʊ.əɾ] ‘bomb’
e: [u] : [ū]  [ˈmu. dú] ‘mute’ : [ˈmʊ. dú] ‘world’
This, however, might be taken either as evidence for the existence of five inherently nasal vowels at the phonological level, or as an epiphenomenon that arises from the possibility of vowels being immediately followed by an abstract nasal coda or not. In other words, under a disegmental analysis, the contrast between [mi.tu] ‘myth’ and [mi.tu] ‘I lie’ (the latter being underlying /miN.tu/) would be of the exact same sort of the contrast between [mi.tu] ‘myth’ and [mi.su] ‘mixed’, where the contrast in question is between two words that are identical except for the absence or presence of the segment /s/ in the coda of the first syllable.

### 2.3. Nasal vowels as disegmental underlying structures

On the other hand, Mattoso Câmara’s (1953, 1970) hypothesis and further elaborations of it (cf. Morais Barbosa 1962, Lemle 1965, Mira Mateus 1975, Almeida 1976, Cagliari 1977, Pardal 1977, Parkinson 1983, Wetzels 1991) is that nasal vowels are underlyingly biphonemic (an oral vowel followed by a nasal coda), involving assimilation of nasality and deletion of the coda segment.5

Under this approach, the five surface nasal vowels [i, e, ì, o, õ] come from underlying forms /iN, eN, aN, oN, uN/, respectively, where N stands for the nasal segment in coda position, whose feature specifications is a matter of debate, and varies among the analyses of the Mattosian family.

\[
(3) \quad \text{Derivation of nasal vowels in Mattoso Câmara (1953, 1970):}
\]

/boN/ underlying representation
bôN effect of regressive nasalization
bô deletion of word-final nasal archiphoneme
[bô] Surface form (written <bom>)

Such an analysis is parallel to Chomsky & Halle (1968)’s postulate that the alternation in the word-final nasal cluster in pairs such as damn/damnation is the result of a nasal deletion rule. Addition of the affixal sequence –ation provides a way for the final /n/ to syllabify in a different syllable than the preceding /m/, and thereby escape deletion. For Mattoso Camara, the existence of pairs such as bom/boníssimo ‘good/very good’, lã/laneiro ‘wool/woolmaker’, fã/fanático ‘fan/fanatic’, and som/sonoro ‘sound/full sounding’ provided evidence for the underlying nasal consonant in words such as bom [bô], lã [lì̂], fã [fã] and som [sõw].

---

5 Moraes & Wetzels (1992) point out that, even before Mattoso Câmara, the biphonemic analysis had been suggested by Trager (1943). However, it was not until Mattoso’s work that it began to be fully developed.

It should be noted, however, that alternations such as *lã/laneiro* are not necessarily informative to the status of an underlying nasal consonant, as epenthetic/linking [n] shows up in word-pairs where there is no evidence for an underlying nasal in the base, such as *Tupi/Tupinista* ‘Tupi/Tupinologist’ and *faraó/faraônico* ‘pharaoh/pharaonic’.

At first blush, pairs like *Tupi/Tupinista* and *faraó/faraônico* might not automatically be taken as robust counter evidence for the idea of an underlying /n/ at the end of the bases. Actually, such pairs (much like *lã/ laneiro, fã/fanático* or *bom/boníssimo*) can be taken as evidence for underlying bases like *tupin* and *faraon*, since the feature specifications of the consonant /n/ that emerges in the complex words (i.e. *Tupinista, faraônico*) is not predictable from the suffix alone or from any morphophonological rule of epenthesis (which tend to give rise to [z] in Brazilian Portuguese). This seems to be in line with the widely accepted analysis by Bisol (1992) whereby all oxytone words ending in an open syllable at the surface would have an underlying consonant at the end, giving rise to a heavy syllable that attracts the stress to itself.\(^6\)

That granted, we still face the question of why such root-final /n/ would nasalize the preceding vowel in some simplex words (e.g. *lã*), but not in others (e.g. *faraô*). By Mattoso’s analysis, all such cases should end with a nasal vowel, contrary to fact. This is an issue for us as well. Would some roots include an underlying /n/ that isn't the source of the nasality of the preceding vowel (which may or may not be nasal after /n/ deletes, depending on the case) while other roots don't include such underlying /n/? Perhaps. And this calls for further investigation. For now, our best bet is that words like *Tupinista, faraônico, fanático* and *lunar* (whose respective ‘bases’ are *Tupi, faraó, fã*, and *lua*) are monomorphemic forms, lexicalized fossils of what used to be a complex word derived from a root ending with /n/, and are clearly not numerous.

In the vast majority of cases, the consonant that emerges between the root and the many possible suffixes is [z] – clearly a default epenthetic consonant in the language – even when the root ends with with a consonant (e.g. *odor* ‘smell’ > *odorzão/*odorão* ‘strong smell’). This is the productive pattern of the language, attested with lots of suffixes (e.g. *-ão, -inho, -ológo, -ologia, -eiro, -aço, -ice, -ista, -ada*, etc) applied to a large number complex words whose simplex counterpart ends with a nasal vowel. For instance, from

---

\(^6\) We are thankful to an anonymous reviewer for calling our attention to this possibility, which prompted us to refine and strengthen our argument. We are also extremely thankful to Vitor Nóbrega, with whom we discussed this matter, and who helped us build a broader empirical base from which our final argument was built.
cupim ‘termite’ [ku.ˈpiʃ], one finds cupinzeiro ‘termite nest’ [ku.piʃ.ˈzej.ɾu], not cupineiro *[ku.piʃ.ˈnej.ɾu]; from marron ‘brown’ [ma.ʃiøw], one finds amarronzado ‘(turned into) brown-ish’ [a.ma.ʃiøw.ˈza.du], not *[a.ma.ʃiøw.ˈna.du]; from maçã ‘apple’ [ma.ˈsɐ], one finds maçãzada ‘a stroke made with an apple’[ma.ʃɐ.ˈza.ʁɐ], not maçãna *[ma.ʃɐ.ˈna.ʁɐ], and so on.

The pattern can be attested in an even clearer way when we observe words formed from the same bases usually brought to make Mattoso’s point. For instance, from bom [bɔm], fã [fɐ] and som [sɔm], one finds bonzão ‘very good’ [bɔ.ˈzɐw], fanzaço ‘super fan’ [fɐ.ˈza.so] and sonzeira ‘impressive musical sound’ [sɔ.ˈzej.ɾa], respectively. The same is true for for words like the ones we brought up (e.g. Tupi, faraô), which don’t exhibit a nasal vowel in the base but do exhibit a surface [n] in some apparently derived words (e.g. Tupinista, faraônico). When such bases combine with most suffixes, we systematically observe an epenthetic [z] between the base and the suffix. For instance, from faraô [fa.ɾa.ˈɔ], we have faraozaço ‘super faraô’ [fa.ɾa.ɔ.ˈza.so], faraozice ‘the set of properties of being a pharaoh’ [fa.ɾa.ɔ.ˈzi.si], and enfaraozado ‘turned into a pharaoh’ [ɛɾ.ɾa.ɾa.ɔ.ˈza.ɾu].

Given all these additional facts, we can safely conclude that Mattoso’s examples of apparently complex words with an epenthetic [n] are not evidence for the source of nasality of vowels being an underlying consonant, as the two the presence of [n] in complex words and the nasality of the last vowel in simplex words may or may not exist independently of each other.

2.4. Open questions

As we discussed above, arguments that definitively decide between the monosegmental vs. disegmental analysis of nasal vowels can be somewhat hard to come by. For both analyses, what seems at first to be evidence in favor of that approach can be interpreted in different terms.

While simplicity of the underlying inventory favors the disegmental analysis (seven underlying vowels instead of twelve), arguments against abstraction favor the monosegmental analysis. In the remainder of this paper, we aim to break out of this analytic stalemate by introducing experimental results from invented language games.

---

7 These last three forms are neologisms, not common words. This in fact strengthens the point being made, since it reveals the internal grammatical mechanisms behind the intuition of the native speaker when forms are not lexicalized one way or another. Definitely, the novel forms just presented are much more natural than their counterparts with [n] in place of [z].
Before that, however, we must present further details about the subinventory of nasal vowels and their realization.

2.4.1. On the proper selection of the nasal subset

As noted above, the set of five nasal vowels [I, ē, û, ŭ, ŭ] is a proper subset of the set of seven oral vowels [i, e, ɐ, o, u], except for the positive specification for the [nasal] feature. This raises the question of how that particular subset is ‘extracted’, excluding mid-open vowels, instead of any other logical possibility.

Under the monosegmental approach, there seems at first to be no way other than simply defining that subset as given. While one might be tempted to say that the reduced inventory of nasal vowels is somehow the less marked of the 7 oral vowels, in fact it is not at all obvious that the closed-mid vowels [e, o] are less marked than the open-mid vowels [ɛ, ɔ] in Brazilian Portuguese, as dialects of the Northeast prefer the latter pair as unmarked defaults (see Nevins 2012 for discussion). Choosing these 5 vowels as a subset without context-free markedness in hand therefore might lead one to speculate that the subinventory of nasal vowels arose diachronically (as well as synchronically, if we think in terms of acquisition) as a selection of the vowels which specifically are the most compatible with the [+nasal] specification, in terms of perception or the contrast between [-nasal] and [+nasal]. In principle, high and high-mid nasal vowels (i.e. [I, ē, ŭ, ŭ]) would be favored over low-mid and low nasal vowels (i.e. [ɛ, ɐ, ɔ]), as the former allow for more resonance of the nasality than the latter. This would suggest a potential explanation for why [ɐ] and [ɔ] are not ‘selected’, but it remains mysterious why [û] is. Although it is true that [û] is a higher ver-

---

8 According to Ruhlen (1975)’s study of universals of nasal vowel inventories, no language has more nasal vowels than oral vowels in its inventory.

9 While the description of the nasal vowels as neutralizing the contrast between open-mid and close-mid vowels [e, ɐ] and [o, ɔ] is virtually unanimous in descriptions of Brazilian Portuguese phonology, there happens to be at least one dialect exhibiting a nasal version of the open-mid vowel [ɔ]. We have come across a native speaker from Paraná who pronounces nasal [ɔ] in very restricted context, the overarching generalization of which is still elusive. At a first approximation, she pronounces [ɔ] in stressed syllables of paroxytonic words whose last syllable is [m]–[me], with some restrictions that seem to be partially related to the category (i.e. whether the word is a verb, a noun, a deverbal noun, etc), and partially related to lexical idiosyncrasies. She does not, however, mirror this pattern with [û], which is absent in her dialect, for reasons irrelevant to the present discussion. While such dialects require further study – and potential merit an article in and of themselves, for present purposes it is is important to emphasize that wherever she has [ɔ] there is no possible minimal pair with words that have [ɐ] or [o] in place of [ɔ], as far as we know. In other words, this dialect does not show six contrasting nasal vowels, as [ɔ] is an conditioned variant of [o].
sion of [â], we still cannot easily make a clear cut between [i, ë, ŋ, ū] and [ē, ê, ṭ] in terms of features responsible for height, since existing [ë] is in fact a little lower than non existing [ê] and [ū].

Under the disegmental approach, the same problem arises, albeit from a different angle. If nasal vowels really are the result of an oral vowel assimilating the [+nasal] specification of an immediately following abstract coda, then it is not immediately clear why there should be a phonotactic constraint preventing only [ɛ] and [ɔ] from combining with the abstract nasal coda. This seems as stipulative as simply defining the set of nasal vowels as being [i, ê, ë, ŋ, ū]. Notice, though, that the phonotactic restriction preventing [ɛ] and [ɔ] from combining with the abstract nasal coda could be easily defined in terms of feature specifications that single out the mid-open vowels [ɛ] and [ɔ], leaving [a] aside. Perhaps a good way of ‘naturalizing’ this phonotactic restriction would be to work out the fine details of the feature specification of the abstract nasal coda, so that we could blame its phonetic properties for why only [ɛ] and [ɔ] are not combinable with it. This is roughly what we will suggest in the remainder of this paper: the second element in the disegmental sequence that composes nasal vowels is a [+high] glide, [j] or [w], that subsequently raises the nucleus with which it combines. Anticipating the result, we will claim that, as opposed to the other four nasal vowels, the surface form [ð] is not the result of an underlying structure /aN/ plus assimilation of the [+nasal] feature, plus coda deletion. Rather, we claim that surface [ð] comes from an entirely distinct inherently nasal vowel, at the phonological level. Section 4 elaborates on this, mostly on the basis of evidence presented in section 3.

2.4.2. On the proper phonetetic characterization of nasal vowels

When we presented the minimal pairs in (1) above, we warned the reader that those transcriptions were somewhat simplified. A more accurate transcription would be as in (4).

(4) Contrast between nasal vowels (accurate transcription)

a: [i] : [ê]  
[si] ‘yes’ : [se] ‘without’

b: [i] : [ð]  
[fj] ‘end’ : [fð] ‘fan’

c: [i] : [ô]  
[küp] ‘termite’ : [küpô] ‘cupon’

d: [i] : [û]  
[fû] ‘kidney’ : [fûw] ‘rum’

e: [ê] : [ð]  
[se] ‘without’ : [sð] ‘sane(fem)’

An important aspect of the nasal versions of [e] and [o] is that, in many contexts, they are pronounced as diphthongs, with a glide that is either [high, front] or [high, back] depending on the articulatory proximity from the vowel’s target position. This glide is particularly salient for nasal [e] in word-final position. For nasal [o], it is also the case that it has a strong tendency to be followed by a glide in word-final position, although, in some cases, this may be somewhat hypoarticulated, giving rise to something akin to [õ]. Word-initially and word-internally, the appearance of these glides is more variable and more idiosyncratic of certain dialects, but it is still possible (cf. Abaurre & Sandalo (2009)). Crucially, in all dialects and all phonotactic positions, there is no possible phonological contrast between [ẽ] and [ẽ], or between [õ̃] and [õ]. We therefore take the latter, surface phonetic realization of these nasal sequences to be largely identical to their underlying phonological specification, a conclusion anticipated in the work of Parkinson (1983).

As for the nasal versions of [i] and [u], the accuracy of the transcriptions in (4) might be questioned. However, as shown by Moraes & Wetzels (1992), there is robust experimental evidence that the nasal versions of [i] and [u] are consistently longer than oral [i] and [u], respectively, which gives support to the transcriptions in (4) above. As a matter of fact, Moraes & Wetzels show that all nasal vowels are longer than their oral counterparts across the board, which would then be evidence for a dissegmental analysis, under the standard assumption the nasal coda, even if further deleted, contributes with a mora to the nasalized vowel. In Section 4, we propose is that the nasal coda is a glide (a conclusion similar to that of Parkinson 1983, who discusses, based on Lacerda, the fact that the locus of nasality was in the second half of nasal vowels).10 This explanation ‘hugs the phonetic ground’ quite closely, explaining a range of facts, including the phonetic

10 Parkinson (1983) in fact proposed that nasal vowels originate not from oral vowel + nasal glide, but oral vowel + nasal vowel. As the realization of this element shows the shorter duration and greater constriction typical of a vowel (and generally does not contribute to syllable count), we characterize it as a glide. Moreover, we depart from Parkinson’s proposal in treating the nasality of [θ] as not due to a glide.
realizations of nasal vowels as diphthongs, their extra lengthening, their co-occurrence with other codas, as well as the different phonological patterning of [ë] compared to the other nasal vowels, shown in sections 3 and 5. Notice that, among all five nasal vowels, [ë] is the only one that is never realized as a diphthong, suggesting that it should be treated separately, as a single underlyingly nasal phoneme.

3. Invented transformational language games

In order to test hypotheses about the underlying phonotactic structure behind nasal vowels, we invented two complementary types of language games, both of which involved some kind of transformation affecting the codas of every syllable of the words. As such, we followed Ohala (1986)’s dictum, in his “Consumer’s Guide to Evidence in Phonology”, that invented language games are some of the most revealing ways to explore phonotactic rules and representations.

3.1. The o-replacement game

The first of these games is we call here the O-Replacement Game (informants were not explicitly taught a name for this game). It involves replacement of all of the vowels in a word by the vowel [o]. Thus, *Universidade Federal do Paraná* becomes *Onovorsodôdo Fodorôl do Poronô*. Participants trained to fluency on this game for about 45 minutes, without exposure to nasal vowels, once tested on nasal vowels, will display results that crucially arbitrate between the monosegmental and disegmental analyses, as shown in (5):

(5) Predictions for the ‘translation’ of [a.le.’křňj] into the O-Replacement Game:

-monosegmental analysis: [a.le.’křňj] → [o.lo.’kro]
-disegmental analysis: [a.le.’křňj] → [o.lo.’krōš]

In the monosegmental analysis, replacement of the word-final nasal

---

11 The data we present in this section comes from a piece of a larger research project that we both conducted with Adelaide Silva (UFPR) and student research assistants (Jennifer Albuquerque, Camila Bozzo, Felipe Clemente, Gustavo Nishida, and Gesoel Mendes), between 2006 and 2008. Results from a pilot pair of experiments that gave rise to the experiment reported here have been previously presented at the *V Congresso Internacional da Associação Brasileira de Linguística*, held in Belo Horizonte, in 2007 (cf. Guimarães, Bozzo & Mendes 2007).
vowel /ũj/ should induce replacement just like any other vowel – since, according to this hypothesis, there are twelve underlying vowels, replacing one should be just like replacing any other. By contrast, in the disegmen-
tal analysis, the word-final nasal in [a.le.'krįj] <alecrim> is the result of a sequence of nasal vowel plus nasal consonant, and hence, [o.lo.'krōw] should induce the same process of nasalization-followed-by-deletion il-
lustrated in (3), yielding a nasalized final vowel. (This assumes that the game, a morphological process, effectuates replacement prior to either of the processes in (3)).

Ten participants at Universidade Federal do Paraná, Brazil, were trained for one hour to play this game, without any nasal-vowelled stimuli in the input. The input consisted of oxytonic (stress-final) words of either one, two, or three syllables in length. In the test session, each participant was asked to transform 110 words, 45 of which had final nasal vowels (9 from each category), randomly intermixed with 65 fillers. Each informant did this three times, totalling 330 responses, of which 135 corresponded to target words containing nasal vowels in the input. Responses were recorded, and three transcribers subsequently judged whether the target vowel was nasal or not. The few transcriber disagreements were resolved by spectro-
graphic analysis. In Table 1 below, the results of nasalization in the output for words like [a.le.'krįj] → [o.lo.'krōw] are shown in terms of percentages, summed over the 10 participants:

<table>
<thead>
<tr>
<th>nasal vowel in the input</th>
<th>% of realization as [ow]</th>
<th>% of realization as [o]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ũj]</td>
<td>99.6%</td>
<td>0.4%</td>
</tr>
<tr>
<td>[ẽj]</td>
<td>99.6%</td>
<td>0.4%</td>
</tr>
<tr>
<td>[ow]</td>
<td>98.5%</td>
<td>1.5%</td>
</tr>
<tr>
<td>[ũw]</td>
<td>96.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td>[ẽ]</td>
<td>34.8%</td>
<td>65.2%</td>
</tr>
</tbody>
</table>

Table 1: Percentages of nasalization in the O-Replacement Game for nasal-vowel inputs

With the exception of the last row (to which we will return), these results strongly support the disegmental analysis of nasal vowels. There would be no reason that nasalization should be retained in the output of the O-Replacement Game under the hypothesis that nasal vowels are under-
yling, atomic segments.12

12 In the discussion below, we consider these results in light of an autosegmental hypothesis for nasal vowels, in which nasality can dissociate, float, and reassociate independently of segmental hosts. Such an analysis will also be shown to be incompatible with our combined set of experimental results.
3.2. The coda-deletion game

The same ten participants who participated in the experiment described in the previous section also participated in another experiment, carried out with an entirely parallel methodology. For the second experiment, we invented a game called the Coda-Deletion Game. Again, the informants were trained for about 45 minutes to play this game, without any nasal-vowelled stimuli in the input. The input consisted of oxytonic (stress-final) words of either one, two, or three syllables in length. In the test session, each participant was asked to transform 110 words, 45 of which had final nasal vowels (9 from each category), randomly intermixed with 65 fillers. Each informant did this three times, totalling 330 responses, of which 135 corresponded to target words containing nasal vowels in the input. Responses were recorded, and three transcribers subsequently judged whether the target vowel was nasal or not. Transcriber disagreement was resolved by spectrographic analysis.

The Coda-Deletion Game works as follows: all coda consonants must be deleted, resulting in only (C)V syllables, and thus *car.tei.ros e.pe.tos* becomes *ca.te.ro e.pe.to*. For nasal-vowelled stimuli, again the monosegmental and disegmental hypotheses diverge, as indicated in (6):

\[
\begin{align*}
\text{monosegmental analysis:} & \quad [\text{a.le.krʃj}] \rightarrow [\text{a.le.krʃj}] \\
\text{disegmental analysis:} & \quad [\text{a.le.krʃj}] \rightarrow [\text{a.le.kri}]
\end{align*}
\]

According to the monosegmental analysis under which nasal vowels such as *[i]* are inherently nasal, there should be no reason for the Coda-Deletion Game to affect the nasality of this vowel. On the other hand, according to the disegmental analysis, deletion of the final nasal consonant will bleed nasalization. (This assumes that the game, a ‘morphological’ alteration, applies prior to either of the processes illustrated in (1)). In Table 2, the results of nasalization in the output for words like *[a.le.krʃj] → [a.le.kri]* are shown in terms of percentages, summed over the ten participants:

\[^{13}\text{Some informants played the O-Replacement Game before the Coda-Deletion game, while other played the games in the reverse order.}\]

nasal vowel in the input | % of realization as [ȷw] | % of realization as [o]
---|---|---
[ɪj] | 1.1% | 98.9%
[ɛj] | 1.2% | 98.8%
[ȷw] | 11.4% | 88.6%
[ʊw] | 3.7% | 96.3%
[ɵ] | 99.2% | 0.8%

Table 2: Percentages of nasalization in the Coda-Deletion Game for nasal-vowel inputs

Again, with the exception of the last row (to which we will return), these results strongly support the *dissegmental* analysis of nasal vowels: nasality on the vowel is lost as the result of a rule of coda deletion. There would be no reason that nasalization should be lost in the output of the Coda-Deletion Game under the hypothesis that nasal vowels are underlying, atomic segments.

### 3.3. Partial conclusions from the results of both experiments

To summarize the result of both games in tandem: In the Coda-Deletion Game, participants did not keep the nasality of putative nasal vowels. This makes sense if the source of nasalization is the (deleted) archiphoneme /N/ of Mattoso Câmara (or some more fully specified nasal segment in coda position). In the O-Replacement Game, participants did not replace a putative nasal vowel with [o], but rather with [ȷ]. This makes sense if the /N/ was still around to cause nasalization. Without assuming a disegmental underlying structure, these results are surprising, because in the O-Replacement Game, the informants did not follow instructions “to the letter”, but actually replaced the vowel with [ȷ] instead! Similarly, in the Coda-Deletion Game, the informants did not follow instructions “to the letter”, but actually mucked with the surface nasality of a vowel. Taken in tandem, then, (and for the moment, still ignoring the results with input [ɵ]), these results provide more evidence for the disegmental analysis than for the monosegmental analysis.

Note that indeed a third logical possiblile is an *autosegmental* analysis, under which the nasality is neither associated with an underlying nasal vowel nor with an underlying nasal consonant, but rather exists in the representation as an independent, autonomous segment (possibly as a ‘prosodic’ property of the syllable as a whole) with the potential to dock, disassociate, float, and reassociate with segmental hosts (cf. Hall 1943). However, the results of the Coda-Deletion Game do not support an autoseg-
mental analysis, since the nasal property of these syllables is not kept when a consonant is deleted, thereby squarely placing the locus of nasality as an inherent property that comes and goes with as part of a specific consonant.

3.4. Control for orthographic interference

In what follows, we turn to a discussion of the divergent patterning of [ɐ], which displays an apparently monosegmental character, as opposed to the other vowels. Before turning to these conclusions, however, it is important that we clarify one aspect of the experiments described above. All target words of that experiment were oxytonic. The most important reason for this choice of design is that it allowed for a uniform comparison of nasal vowels in monosyllabic, disyllabic, and trisyllabic words: all of them contained the nasal vowel in a final stressed syllable (e.g. [fiu̯], [mu.tu̯], [u.ru.ku̯]). As Brazilian Portuguese virtually does not allow unstressed final nasal vowels (pending a few exceptions such as [ɔh.'fɐ̃̃o̯] ‘orphan’), there would not have been sufficient numbers of stimuli to create balanced materials. Final stress thus allowed for complete control over stress placement across all three syllable lengths.

However, an unavoidable consequence of final-stressed nasal vowels is that by the rules of orthography, nasal vowels in stressed syllables of oxytonic words are not all spelled the same way (as opposed to proparoxytonic and paroxytonic words, where all nasal vowels in stressed position are spelled with a following letter corresponding to a nasal consonant: <m> or <n>). That is, in oxytonic words, nasal vowels [i, ê, ô, ü] are spelled as <im, em, om, um>, respectively, whereas [ɐ] is spelled as <ã>.

At first blush, this feature of the experimental design might seem to be detrimental to the reliability of the results, given that the different behavior between [ɐ] and the other nasal vowels under the two language games could be attributed to some sort of orthographic bias, as if the informant were writing up the transformed version of the word in an imaginary blackboard, and then reading it aloud. When we first designed the experiment,

14 For instance, given the words [fiθ] ‘frog’ and [fiũ] ‘rum’, spelled as <ra> and <rum>, the informant could write up their transformed versions in his/her ‘imaginary blackboard’ as <ro> and <rom> (for the O-Replacement Game) and <ra> and <ru> (for the Coda-Deletion Game), and then read them aloud, leading to the attested pattern through orthography, skipping phonology. This is a logical possibility which will be shortly dismissed on the basis of further evidence. Notice, however, that replacing <ã> with <o> under the is not the only logical possibility for the O-Replacement Game. The informant could just as well treat <ã> as a complex orthographic chunk, made up of <a> and the diacritic <~>, and then replace <a> with <o>, leaving <~> in place, resulting in <Ô>, which should be read aloud as [ʊ].

we weren’t necessarily expecting one of the nasals to be monosegmental alongside the others being disegmental. If, for both games, the informants treated all five nasal vowels equally, we would have had evidence that they weren’t behaving according to any orthographic bias.

As we now know, the informants had systematically treated [B] differently. Does this mean that, perhaps as the result of experimental design, our informants were all ‘reading letters from an imaginary blackboard’? Not necessarily so. First of all, as a matter of logic, the special behavior of [B] could be a consequence of either an orthographic bias or an actual phonotactic difference between [B] and the other nasal vowels. Thus, in the worst case, the results would be inconclusive. However, we have good reason to believe that the amount of orthographic bias was negligible in both experiments. We say so because we carefully include, both in the training session and in the experiment itself, some filler words that are very telling of whether the informant is speaking “by the letter” or “from the phonology”.

Given the general orthographic conventions of Brazilian Portuguese, letters <c> and <g> correspond respectively to [s] and [z] when they appear immediately before <i> or <e>; and correspond respectively to [k] and [g] elsewhere. We used this as a control for identifying orthographically-biased informants during the training session, and discarding them. In (9) below, we present some of the control words displayed in the following way. In the first column we provide the ordinary pronunciation of the word followed by an English gloss. In the second column, we provide the ordinary orthographic representation of it. In the third and fourth columns, respectively, we have the predicted pronunciations of those words under the O-Replacement Game with and without orthographic bias. Crucially, all our ten informants systematically pronounced the output form of the game as in the last column.

(7) **Control Words**

<table>
<thead>
<tr>
<th>Ordinary Speech</th>
<th>Spelling</th>
<th>Reading Letters</th>
<th>Realizing Phonemes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[si.ka.ˈtris]</td>
<td>&lt;cicatriz&gt;</td>
<td>[ko.ko.ˈtrɔs]</td>
<td>[so.ko.ˈtrɔs]</td>
</tr>
<tr>
<td>[ˈsiz.me]</td>
<td>&lt;cisma&gt;</td>
<td>[ˈkoz.mo]</td>
<td>[ˈsoz.mo]</td>
</tr>
<tr>
<td>[ˈsi.ɾa.ˈsɔw]</td>
<td>&lt;girassol&gt;</td>
<td>[ˈɡo.ɾo.ˈsɔw]</td>
<td>[ˈɡo.ɾo.ˈsɔw]</td>
</tr>
<tr>
<td>[ˈziʃ]</td>
<td>&lt;giz&gt;</td>
<td>[ɡos]</td>
<td>[ɡos]</td>
</tr>
</tbody>
</table>

These controls for orthographic bias suggest that our speakers’ divergent outputs for nasal [B] as opposed to the other vowels was not due to a task-
related effect involving spelling, but rather reflects a distinct pattern within the phonology of these vowels themselves. We present our interpretation in the next section.

4. The divergent patterning of nasal [ɐ]

Recall that in both invented language games reported above, the results implicate the presence of a nasal coda element, seemingly the Mattosian archiphoneme /N/. However, the phonetic reality of the realization of nasal vowels, as well as a number of facts to be discussed below, suggest that while there is a nasal coda in the sequences in question, it is a glide (in fact, two glides, either /w/ or /j/), and not a nasal obstruent.15

Assume all of the ten combinations of two nasal glides with five vowels. The homorganic combinations [öw, ūw, ĕj, ĭj], as words like [sōw] <som> ‘sound’ and [ŏ.mĕj] <homem> ‘man’ are typically pronounced as [sōw] and [ŏ.mĕj] – in fact, as Parkinson (1983) points out, when words like homem monopthongize, they also denasalize, yielding [Ŏ.mi]. This makes sense if the locus of nasality and the locus of the glide is one and the same. The combinations [ŭj, ĕw, ĭw] do not exist underlingly (though may result as reductions of disyllabic words like [pĭ.ũo] ‘pine’ [fû.ĭ] ‘bad’) – a restriction on combinations of high vowels and high glides with conflicting specification for color (palatality/labiality) that is crosslinguistically recurrent.

The combination [ŏj] exists as a morphologically derived form (plural of nouns ending in <ão> [ŏw], and the inflected form of the verb por ‘put’), but not as an underived noun that forms minimal pairs with anything. That leaves only two combinations: [ēw] and [ĕj], and these two are indeed underlingly contrastive. In other words, a phonetically realized nasal glide in the coda of a syllable is only ever potentially contrastive with a ‘purely’ nasal vowel when the nucleus is /a/. Nasality in nasal vowels, by hypothesis, always comes from a nasal glide, except in the case of [ɐ], where it cannot come from a glide, as the glides are reserved for the contrastive nasal diphthongs [ĕj] and [ēw].

15 One might be tempted to reduce the inventory of underlying nasal glides from two to one, say a [-cons,-voc,+nasal] element that is underspecified as to whether it is a labial and palatal glide, and thus predictably homorganic with a preceding nucleus. It is not clear at present how such a representation would account for the independently needed presence of nasal palatal and nasal labial glides with /a/ as a nucleus (e.g. /ĕj/ and /ěw/), although we leave such a possibility open for future research.
Consider the minimal triplet [mɐ̃, mã, mão] <mam, mãe, mão>. A nasal glide is simply not possible as a source of nasalization for [ɐ] without obliterating the contrast between it and these other two diphthongs. The nasality of [ɐ] does not come from the glide, but is inherent to this vowel. In fact, as mentioned in Section 2, nasalized [ã] is an entirely inaccurate phonetic transcription. It should not even be written with the same grapheme, as it is much higher. Unlike all other four vowels, the nucleus is significantly raised, to the point of being nearly a central vowel. It is probable that learners (synchronically as well as diachronically) had/have no reason to assume it is related to [a]. Only the orthography suggests that [ɐ] is related to [a], and even then, the fact that it is orthographically represented differently from the four other nasal vowels suggests a reflection within the orthography of the underlying intuition that nasality in [ɐ] is monosegmental, while nasality in [i, e, õ, ô] is disegmental.

With the interpretation of nasality in [ɐ] as monosegmental, all of the experimental results above fall in place. In the O-replacement game, /ã/ is wholly replaced, thereby obliterating its nasality. For the other four vowels, nasality is not a property of the vowel, and hence remains intact under replacement. By contrast in the Coda-Deletion game, nasality, as a property of the glide, is removed for all of the four disegmental sequences. For /ã/, however, it remains intact, as for this vowel alone it is not localized in the coda.

5. Converging evidence from ‘Gualin do Tetecá’

We now turn to independent evidence for the distinct patterning of nasal [ɐ] as compared to the other four nasal vowels. The data come from an existing language game played in the Catete neighborhood of Rio de Janeiro, Gualin do Tetecá (Língua do Catete, with syllable inversion). As such, it largely identical to the rule employed in games such as French Verlan (Lefkowitz 1989).16 Here are some basic examples of how it works in sentential contexts:

---

16 Actually, Gualin do Tetecá is not the only language game based on Brazilian Portuguese which involves reordering of syllables. For instance, Butzen (2005) documented one such language game played around the cities of Joaçaba and Herval d’Oeste, in Santa Catarina state. Recently, the mass media has uncovered another such language game played in the city of Sabino, in São Paulo state (cf. https://www.youtube.com/watch?v=_YIJTbjFauU and http://www.youtube.com/watch?v=Rc62nRyOA24).
As the reader will notice, a number of interesting processes occur, such as the ‘double encoding’ of the infinitive in marfuzar (presumably marfu is not transparent as an infinitive, hence adds the epenthetic z and then the first conjugation infinitive ending). Of even further interest is the fact that words undergoing transformation in the game virtually always end up with final stress (e.g. Catete, with penultimate stress in the original, becomes Tetecá, as such neither preserving the original location of stress nor the originally-stressed syllable of the input). It is somewhat notable that a number of language games in Brazilian Portuguese result in iambic patterns/final stress (including Língua do Pê; Guimarães & Nevins 2012). We speculate that final stress within the context of language games is preferred for a very specific reason: because of its interaction with vowel reduction. There is a very striking asymmetry in the language: post-tonic vowel reduction is virtually obligatory in Brazilian Portuguese, while pre-tonic reduction is optional (see Nevins 2012 for discussion). Language games with antepenultimate or penultimate stress, therefore, would induce neutralizing post-tonic vowel reduction, thereby further obscuring the ‘decoding’ of the game to original underlying inputs in a way that might surpass listeners’ ability to undo vowel reduction and syllable inversion simultaneously.

In any event, of greatest relevance to the current paper is the fact that allophonically nasalized vowels in Gualín do Tetecá show the same asymmetry between /a/ and the other four vowels. Specifically, while allophonically nasalized /a/ preserves its nasalization after syllable-inversion, the other four vowels do not:

(9) a: cama [ˈkɐ̃.mɐ] → macã [ma.ˈkɐ̃] ‘bed’
c: rumo [ˈʁũ.ˌmo] → morrú [mo.ˈxr] ‘route’
d: vinho [ˈvũ.ˌno] → nhoví [ŋo.ˈvi] ‘wine’
e: unha [ˈũ.ˌna] → nhaú [na.ˈu] ‘fingernail’
f: tema [ˈtɐ̃.mɐ] → matê [ma.ˈte] ‘theme’
g: tênis [ˈtɐ̃.níʃ] → nistê [nistê.ˈte] ‘tennis’
h: rima [ˈɾi.ˌmɐ] → marrí [ma.ˈɾi] ‘rhyme’

There is no imaginable way here that an ‘imaginary blackboard’ on which orthographic bias were to play a role would differentiate <cama> and
<tema>, such that nasality should manifest itself only as the result on inversion in the former. These results thereby further argue for a distinct status of nasal [ã].

Now, given that all instances of nasalization in a stressed vowel preceding an immediately following heterosyllabic nasal consonant are predictable, the question arises as to why nasalization should ‘disappear’ in the cases of [i, ê, ô, u], but not [õ], under inversion. We contend that yet again, the nasal glide plays a crucial role here. Specifically, let us assume that nasal [i, ê, ô, u] when occurring without a following glide are unlicensed. (This can be informally captured in a constraint such as *õ#, which is satisfied in the sequence [õw]#; the same phonotactic constraint would hold for the other three vowels in question). As no glide was present in the input /roma/, none will be present in the intermediate output [ma.'xô], and the unlicensed nasalization is removed, yielding [ma.'xo]. By contrast, in the intermediate output [ma.'kõ], no relevant licensing constraints demand the presence of a glide, and this “structure preserving” instance of the sequence [õ]# survives unscathed.

While the specific formulation of the interaction between nasalization, syllable-inversion, and potential denasalization must await further experimentation with speakers of Gualin do Tetecá, the fact that this evidence points towards the same difference between [õ] and the other nasal vowels, specifically in a case where orthography cannot be in any way held responsible, provide converging evidence for our overall hypothesis.

6. Conclusion: Mattoso was 80% correct

We began this paper by contrasting two families of analysis: one in which Brazilian Portuguese has twelve underlying vowels, with that of Mattoso Câmara, in which Brazilian Portuguese has seven underlying vowels, the five nasal vowels being derived from syntagmatic, rather than paradigmatic, contrast. The experimental results in this paper, demonstrating a distinct pattern of results in transformational games of [õ] from the other four nasal vowels, suggests that Mattoso was 80% right: 4 of the 5 nasal vowels are indeed derived from sequences of oral vowel plus nasal consonant. Contrary to Mattoso’s analysis of this nasal consonant as an archiphonemic nasal stop, however, we propose that it is a glide, an analysis that is entirely grounded in the phonetic realizations of these four vowels (Lacerda & Head 1966, Parkinson 1983, Moraes & Wetzels 1992, Abaurre & Sandalo 2009): they frequently wear off glide in their pronunciation on
their sleeves, enjoy a longer duration than corresponding oral vowels, and bear the locus of nasality in the second half of the sequence. As argued in section 2.3 above, Mattoso’s contention that the underlying consonant was a nasal stop, based on alternations such as lâ/laneiro ‘wool/woolmaker’ and fâ/fanático ‘fan/fanatic’ are not informative to the status of an underlying nasal consonant, as epenthetic/linking [n] shows up in morphologically derived nominal/adjectival forms even where there is no robust evidence for an underlying nasal in the base, such as Tupi/Tupinista ‘Tupi/Tupinologist’ and faraó/faraônico ‘pharaoh/pharaonic’. Also, there are pairs where the base ends with a nasal vowel, with the derived word exhibiting [z] instead of [n] between the base and the suffix (e.g. fâ/fanzaço ‘fan/super-fan’) and pairs where the same base (ending with an oral vowel) that forms complex words with an [n] before the suffix (e.g. faraó) also forms complex words with [z] before the suffix (e.g. faraozação ‘super-pharaoh’).

We conclude the paper, however, with an admission that our experimental results, the basis for a hybrid monosegmental/disegmental representation for nasal vowels that is, as far as we know, the first of its kind, as of yet has nothing to say about the nasal vowels found in preconsonantal position, such as panda ‘panda’, tenda ‘tent’, and onda ‘wave’. Note that the offglides possible in words like [bẽ] <bem> ‘well’ and [sõw] <som> ‘sound’ are not obviously arising in words like [fê.de] <renda> ‘income’ and ‘[pô.br] <pomba> dove’. In principle nothing rules out an analysis in which [pê.de] <panda> ‘panda’ and [fiû.br] <rumba> ‘rumba’ literally do involve a nasal consonant, i.e. /panda/ and /hunba/ and not /pânda/ and /huwba/. In this case we should expect no difference between these two words in cases of nasality followed by a heterosyllabic consonant within language games of the sorts we have described here. In fact, Moraes & Wetzel point to suggestive phonetic evidence, based on duration, that vowels of this sort may be distinct from ones in clearly open syllables, such as [fê], and Shosted (2003) documents, using nasal airflow, the existence of a clearly articulated moment of nasal obstruency (a true, albeit brief /n/ in words like panda), whose realization very well may be a cue to learners that the phonological representation actually involves an obstruent (despite the unification between pomba and som suggested by the orthography). Investigating whether Mattoso’s hypothesis is, indeed 100% right for the case of preconsonantal nasal vowel sequences such as panda and pomba, through the use of transformational and autochthonous language games, is clearly the next step for future research in this direction.

17 Though see Abaurre & Sandalo (2009) for some cases where there clearly is an offglide in words of such kinds, restricted to a few dialects.

REFERENCES

HALL, R. The Unit Phonemes of Brazilian Portuguese. Studies in Linguistics 1 (15), pp. 1-6, 1943.


TRAGER, G. *Nota aos artigos de R. Hall Jr. 1943a, 1943b,*. 1943.


*Recebido em: 31/03/2013; Aceito em: 07/06/2013*